

AKZENTWAHRNEHMUNG VON JAPANERN BEI TECHNISCH KONTROLLIERTEN F0-KONTUREN

Akiko Masaki, Miyuki Takasawa und Takayuki Arai

The dominant perceptual cue of accent is said to lie in the fundamental frequency (F0), not only in pitch accent languages like Japanese and Swedish, but also in stress accent languages like English and German. Something not discussed extensively is whether native speakers of different languages use the same cue or whether their perception of accent is influenced by the accentual characteristics of their native language when they perceive accented syllables in another language. In this study we investigate which cues Japanese are more sensitive to when they mark an accented syllable in a two-syllable meaningless word. To compare the results of Japanese subjects with the subjects of other languages we also refer to the results of three non-Japanese. In the stimuli, the F0 contour of both the first and second syllable is controlled. Specifically, the F0-contour of the first syllable varies from falling to rising, while the F0 contour of the second syllable is kept flat. Results show that Japanese subjects are more sensitive to a falling F0 contour than to a rising one. In addition, we find that the perception of accented syllables is categorical, that is, accent is perceived when the first syllable has a falling F0-contour, no matter how steep the declination.*

1. Einleitung

Unter den akustischen Eigenschaften des Akzentes spielt die Grundfrequenz (F0) die größte Rolle, vor allem in Sprachen mit musikalischem Akzent wie Japanisch (Beckman 1986, Sugitou 1990). Bei der Wahrnehmung des Akzentes im Japanischen wird der fallende Punkt der Grundfrequenzkonturen als die einflussreichste Eigenschaft betrachtet (Hayashi et al. 1997). Auch in Sprachen mit dynamischem Akzent, wie Englisch und Deutsch, spielt die Grundfrequenz eine dominierende Rolle bei der Wahrnehmung des Akzentes (Hirst und Di Cristo 1988). Es wurde bisher aber nicht eingehend diskutiert, ob es zwischen Japanern, deren Muttersprache musikalischen Akzent hat, und den Muttersprachlern anderer Sprachen mit dynamischem Akzent bei der Unterscheidung der akzentuierten Silbe in einer fremden Sprache einen Unterschied gibt.

In unserer vorherigen Arbeit (Masaki et al. 2000) haben wir

untersucht, auf welche akustische Eigenschaft Japaner, Deutsche und Spanier reagieren, wenn sie Wörter aus Fremdsprachen hören, die sie nicht kennen. Es ergab sich ein deutlicher Unterschied zwischen Japanern und Deutschen, wenn sie akzentuierte Silben in Wörtern aus dem Norwegischen, Finnischen und Hindi markierten. In diesem Wahrnehmungsexperiment markierten die deutschen Versuchspersonen die Silben als akzentuiert, in denen die Grundfrequenz den höchsten Wert hatte, während die japanischen und spanischen Versuchspersonen die Silben als akzentuiert betrachteten, die eine hohe Anfangsstelle der fallenden Grundfrequenz hatten. Das Ergebnis bei den japanischen Versuchspersonen reflektiert die Wahrnehmungstendenz des Akzentes im Tokyo-Japanischen, in dem die Tonhöhe bei Wörtern mit Akzentkern unmittelbar nach dem Akzentkern plötzlich absinkt.

In dieser Arbeit untersuchen wir noch eingehender den Einfluss sowohl der fallenden als auch der steigenden Grundfrequenzkonturen auf die Wahrnehmung der akzentuierten Silbe durch japanische Versuchspersonen. Die Ergebnisse der japanischen Versuchspersonen werden dann mit denen einer deutschen, einer spanischen und einer polnischen Versuchspersonen verglichen. Da die Daten der europäischen Versuchspersonen gering sind, werden sie nur sekundär zur Ergänzung herangezogen.

2. Wahrnehmungsexperiment

2.1. Stimuli

In unserer vorigen Arbeit war das norwegische Wort "ara" am umstrittensten. In diesem Wort war die Grundfrequenzkontur von [a] in der ersten Silbe deutlich fallend von 250 Hz auf 202Hz, während das [a] in der zweiten Silbe steil von 212Hz auf 315Hz stieg. Die meisten japanischen Versuchspersonen (81%) markierten die erste Silbe als akzentuiert, während die meisten deutschen Versuchspersonen (85.7%) die zweite Silbe als akzentuiert betrachteten.

In diesem Experiment benutzen wir das Wort "ata" statt "ara". Der stimmlose Verschlusslaut [t] wird statt [r] zwischen den ersten und zweiten Vokal eingesetzt, um den Einfluss des Konsonanten auf die Grundfrequenzkontur zu vermeiden. Obwohl die Lautkombination /ata/ in der japanischen Sprache möglich ist, gibt es im Japanischen ein solches Wort nicht.

Der originale Stimulus "ata" wird von einem Japaner aus Tokyo gesprochen. Vor dem Wahrnehmungsexperiment werden die Grundfrequenzkontur des originalen Wortes mittels des digitalen Computers CSL (Kay Elemetrics, 4300B DOS version) technisch variiert und 64 verschiedene Stimuli erzeugt. Die Längen der ersten und zweiten Vokale werden jeweils auf 200ms fixiert, während der Verschlusslaut dazwischen die Länge von 100ms hat.¹ Die Grundfrequenzkontur in der letzten Hälfte der ersten Silbe [a] (100ms) in "ata" wird technisch entweder steigend, fallend oder flach eingestellt.² Die Grundfrequenzkontur der zweiten Silbe wird immer flach mit 8 verschiedenen Tonhöhen (145Hz, 139Hz, 135Hz, 130Hz, 125Hz, 120Hz, 115Hz und 110Hz) eingestellt. (vgl. Abbildung 1) Die ausführlichen Konturen der einzelnen Stimuli werden in den folgenden Abschnitten gruppenweise behandelt.

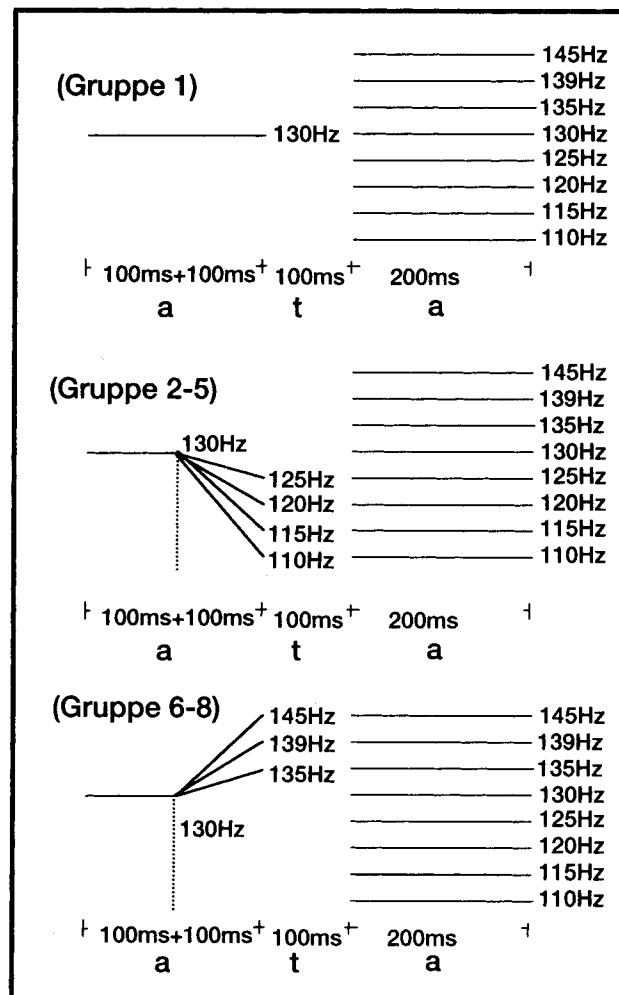


Abbildung 1: Stimuli

2.1.1. Gruppe 1

In Gruppe 1 sind die Grundfrequenzkonturen von "ata" in der ersten und zweiten Silben immer flach. In dieser Gruppe wird die Grundfrequenz des ersten Vokals stets auf 130Hz eingestellt, während die des zweiten Vokals in 8 verschiedenen Tonhöhen variiert wird.

2.1.2. Gruppe 2 - 5

In Gruppe 2 - 5 sind die Grundfrequenzkonturen in der letzten Hälfte (100ms) des ersten Vokals [a] immer fallend. In Gruppe 2 ist die Grundfrequenzkontur leicht fallend, von 130 Hz auf 125Hz. In Gruppe 3 ist der Abstieg etwas steiler, von 130Hz auf 120Hz; in Gruppe 4 von 130Hz auf 115Hz; und in Gruppe 5 fällt die Grundfrequenzkontur sogar von 130Hz auf 110Hz. Auch bei diesen Gruppen sind die Grundfrequenzkonturen der zweiten Silbe immer flach und haben 8 verschiedene Tonhöhen.

2.1.3. Gruppe 6 – 8

In Gruppe 6 – 8 sind die Grundfrequenzkonturen in der letzten Hälfte des ersten Vokals immer steigend: In Gruppe 6 nur leicht, von 130Hz bis 135Hz. In Gruppe 7 ist die Steigung etwas steiler als in Gruppe 6, nämlich von 130Hz bis 139Hz. In Gruppe 8 ist die Steigung am steilsten, nämlich von 130Hz bis 145Hz. Bei diesen Gruppen sind die Grundfrequenzkonturen der zweiten Silbe immer flach und haben 8 verschiedene Tonhöhen.

2.2. Versuchspersonen

Die Versuchspersonen sind 17 japanische Muttersprachler (13 Männer und 4 Frauen), deren Eltern auch Muttersprachler sind. Sie sind alle StudentInnen, die keine speziellen Kenntnisse der Phonetik haben. Sie verstehen Englisch, einige auch andere Sprachen, wie Französisch oder Spanisch, aber sie sprechen diese nicht fließend. Ihr Aufenthalt im Ausland erstreckt sich nicht über ein Jahr.

Auch drei ausländische Versuchspersonen haben an diesem Experiment teilgenommen: einer aus Deutschland, einer aus Spanien und eine aus Polen. Der Deutsche und die Polin sind Architekten und verstehen überhaupt kein Japanisch. Der Spanier ist Student, der seit 5 Monaten in Japan wohnt und als Anfänger ein bisschen Japanisch spricht.

2.3. Versuchsablauf

Das Experiment wurde im Phonetischen Laboratorium der Sophia-Universität durchgeführt. Eine oder zwei Versuchspersonen nahmen gleichzeitig am Experiment teil. Jede Versuchsperson benutzte Kopfhörer.

64 verschiedene Stimuli kommen in einem Experiment dreimal vor und werden nach Zufall zusammengestellt. Diese Stimuli werden durch den Computer (Toshiba DB65P/4RA) zweimal vorgespielt.

Zuerst werden den Versuchspersonen die Beispielwörter ausgehändigt. Auf dem Fragebogen werden die nummerierten Wörter aufgeschrieben. Die Versuchspersonen werden dann aufgefordert, nach dem Vernehmen des Stimuluswortes die akzentuierten Silbe, nämlich entweder "a" oder "ta", zu markieren. Sie dürfen nur eine Möglichkeit auswählen. Wir klicken die Maus des Computers, um jeden Stimulus vorzuspielen. Das Intervall zwischen den Stimuli hängt von der jeweiligen Versuchsperson ab. Das Experiment dauert etwa 25 Minuten.

2.4. Ergebnisse

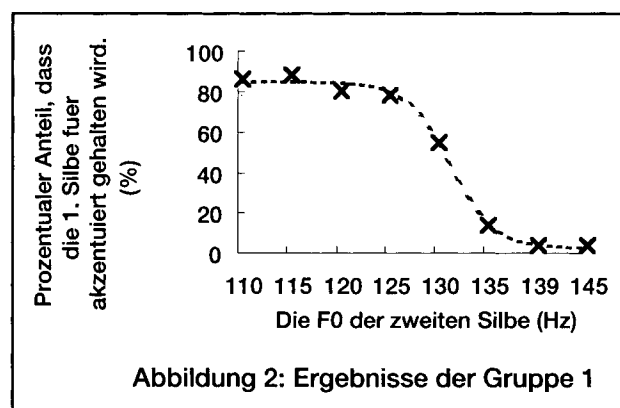
Die durchschnittlichen prozentualen Anteile, bei denen die erste Silbe als akzentuiert wahrgenommen wird, werden gruppenweise in Abbildungen 2 – 9 dargestellt.

2.4.1. Japanische Versuchspersonen

Die Zahl der Antworten für einen Stimulus beträgt insgesamt 51, weil die 17 japanischen Versuchspersonen jeweils dreimal auf einen Stimulus geantwortet haben. Die Ergebnisse werden nach der Grundfrequenzkontur der ersten Silbe (flach, fallend oder steigend) gruppenweise diskutiert.

2.4.1.1. Gruppe 1

Wie in Abbildung 2 dargestellt wird, können die japanischen Versuchspersonen nicht unterscheiden, welche Silbe akzentuiert ist, wenn die Grundfrequenzkonturen bei der ersten und zweiten Silbe flach sind und die erste und die zweite Silbe dieselbe Grundfrequenz, nämlich 130Hz, haben. Wenn aber die Grundfrequenz der zweiten Silbe um 5Hz höher oder niedriger als bei der ersten Silbe ist, halten die meisten Versuchspersonen (etwa 80%) die Silbe, die die höhere Grundfrequenz hat, für akzentuiert.

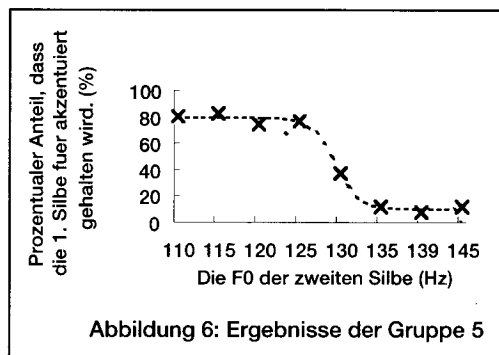
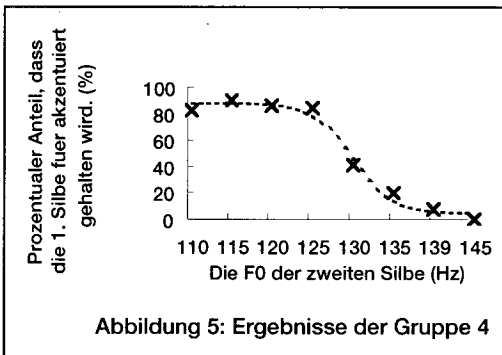
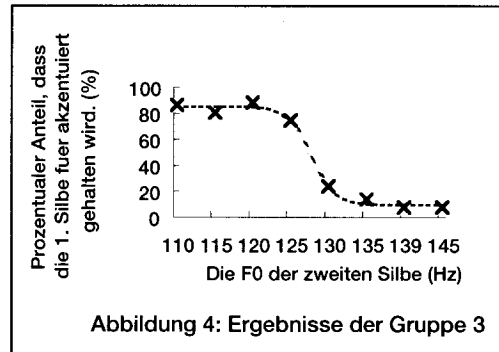
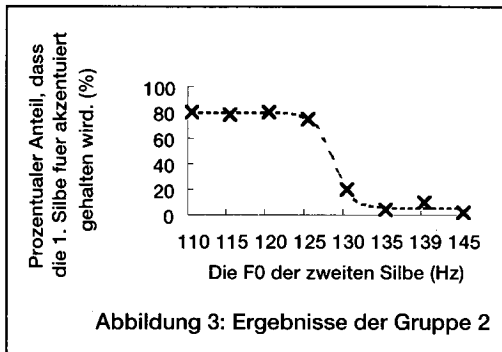


2.4.1.2. Gruppe 2 – 5

Bei den Stimuli der Gruppen 2 – 5 sind die Grundfrequenzkonturen der ersten Silbe in der letzten Hälfte (100ms) des ersten Vokals [a] fallend, während sie in der zweiten Silbe flach sind. Im Mittelpunkt des Interesses steht der Einfluss der fallenden Grundfrequenzkontur und ihrer Steilheit auf die Wahrnehmung des Akzentes.

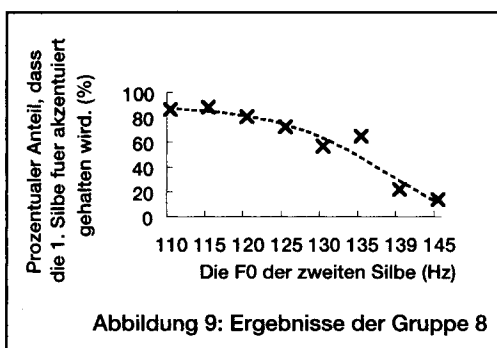
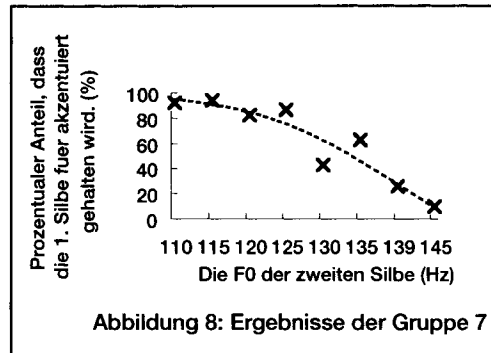
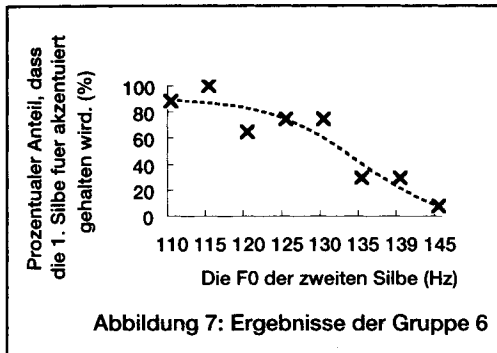
Wenn die Grundfrequenz der zweiten Silbe 130Hz oder höher als

130Hz ist, halten die Versuchspersonen die zweite Silbe für akzentuiert. Dabei scheint die Steilheit der fallenden Grundfrequenzkontur in der ersten Silbe nicht berücksichtigt zu werden (vgl. Abbildung 3 – 6). Wie in Abbildungen 3 – 6 gezeigt wird, wird die akzentuierte Silbe in Gruppe 2 – 5 kategorial³ wahrgenommen.



2.4.1.3. Gruppe 6 – 8

Bei den Stimuli der Gruppen 6 – 8 sind die Grundfrequenzkonturen der ersten Silbe steigend, während die der zweiten Silbe flach sind. Wie in Abbildung 7 – 9 dargestellt wird, unterscheiden sich die Ergebnisse der Gruppe 6 – 8 von denen der Gruppe 2 – 5. Wenn die Grundfrequenzkontur der ersten Silbe von 130Hz auf 135Hz steigt (Gruppe 6), halten 70% der Versuchspersonen die zweite Silbe, in der die Grundfrequenz 135Hz oder höher ist, für akzentuiert. Wenn die Grundfrequenzkontur der ersten Silbe von 130Hz auf 139Hz und von 130Hz auf 145Hz steigt, halten etwa 63% der Versuchspersonen die zweite Silbe, in der die Grundfrequenz 139Hz oder höher ist, für akzentuiert. Wenn die Grundfrequenzkonturen der ersten Silbe steigend sind, unterscheiden die Versuchspersonen die akzentuierten Silben nicht so deutlich, wie in den Fällen, in denen die Grundfrequenzkontur der ersten Silbe fallend ist.



2.4.2. Europäische Versuchspersonen

Da die drei europäischen Versuchspersonen drei verschiedene Muttersprachler sind, dürfen ihre Ergebnisse nicht einfach mit den Ergebnissen der japanischen Versuchspersonen verglichen werden. Obwohl die Daten nur zur Information behandelt und diskutiert werden sollen, werden die folgenden Punkte bemerkt: wenn die Grundfrequenzkonturen in der ersten und zweiten Silbe flach sind (Gruppe 1), gibt es keinen großen Unterschied zwischen Japanern und anderen Muttersprachlern. In Gruppe 2 – 5, in denen die Grundfrequenzkontur in der ersten Silbe fallend ist, wird keine einheitliche Tendenz unter den europäischen Versuchspersonen beobachtet. Wenn die erste Silbe eine steigende Grundfrequenzkontur hat (Gruppe 6 – 8), halten der Deutsche und die Polin die zweite Silbe erst dann für akzentuiert, wenn die Grundfrequenz der zweiten Silbe 145Hz beträgt. Diese Grundfrequenz ist höher als bei den meisten japanischen Versuchspersonen, die die zweite Silbe schon bei einer Grundfrequenz von 135Hz für akzentuiert halten. Um noch eingehender zu diskutieren, werden weitere Daten benötigt.

3. Diskussion

Aus den Ergebnissen des Wahrnehmungsexperimentes können folgende fünf Punkte hervorgehoben werden:

1. Wenn die Grundfrequenzkontur flach ist, übt der Unterschied der Grundfrequenzen zwischen der ersten und der zweiten Silbe einen Einfluss auf die Wahrnehmung des Akzentes aus. Die Versuchspersonen können nicht unterscheiden, ob die erste oder die zweite Silbe akzentuiert ist, wenn die Grundfrequenzkontur flach ist und die Grundfrequenzen der ersten und der zweiten Silbe gleich sind.
2. Wenn die gesamte Grundfrequenzkontur des Wortes "ata" nur fallend ist, halten die Versuchspersonen die erste Silbe für akzentuiert. Auch wenn die zweite Silbe etwas höher (etwa 5 – 15 Hz) als die Grundfrequenz am Ende der ersten Silbe ist (Gruppe 2 – 5), halten die Versuchspersonen die erste Silbe für akzentuiert. Dabei ist aber die Grundfrequenz der zweiten Silbe stets niedriger als 130Hz. Es kann zwar wichtig sein, ob die erste Silbe fallend ist oder nicht, aber die Steilheit der fallenden Grundfrequenzkontur scheint keine Rolle zu spielen, wenn die Grundfrequenz der zweiten Silbe niedriger als die Anfangsgrundfrequenz der ersten Silbe, nämlich 130Hz, ist. Diese Anfangsgrundfrequenz von 130Hz wird sogar von den japanischen Versuchspersonen für den Vergleich mit der Tonhöhe der zweiten Silbe für wichtig gehalten.
3. Wenn man die Ergebnisse in Beziehung zu der Grundfrequenz am Ende der ersten Silbe und der Grundfrequenz der zweiten Silbe setzt, verschiebt sich der kritische Punkt, an dem die erste Silbe von mehreren Versuchspersonen für akzentuiert gehalten wird. (vgl. Tabelle 1) Es kommt darauf an, ob die Grundfrequenzkontur der ersten Silbe fallend oder steigend ist. Wenn die Grundfrequenzkontur der ersten Silbe fallend ist, beträgt der kritische Punkt der Grundfrequenz in der zweiten Silbe 130Hz. Die japanischen Versuchspersonen halten nämlich die zweite Silbe für akzentuiert, wenn die Grundfrequenz 130Hz oder höher ist. Wenn die Grundfrequenzkontur der ersten Silbe steigend ist, liegt der kritische Punkt der Grundfrequenz in der zweiten Silbe zwischen 135Hz und 139Hz. Es kommt daher, dass für japanische Versuchspersonen die Wahrnehmung des Akzentes von den fallenden und steigenden Grundfrequenzkonturen der ersten Silbe abhängt.

		Die FO der zweiten Silbe (Hz)							
		110	115	120	125	130	135	139	145
Die letzte FO der ersten Silbe (Hz)	145	a	a	a	a	a	a		
	139	a	a	a	a	a	a		
	135	a	a	a	a	a			
	130	a	a	a	a	*			
	125	a	a	a	a				
	120	a	a	a	a				
	115	a	a	a	a				
110	a	a	a	a					

Tabelle 1: Die Matrix des kritischen Punktes, wo die erste Silbe "a" fuer akzentuiert gehalten wird.

4. Wie die Abbildungen 2 – 9 zeigen, scheinen die Wahrnehmungseigenschaften der steigenden Grundfrequenzkonturen der ersten Silbe im Vergleich zu den fallenden Grundfrequenzkonturen nicht kategorial zu sein. Die relativ abrupt fallenden Konturen in den Abbildungen 3 – 6 zeigen eine kategoriale Wahrnehmung, während sich die Konturen in den Abbildungen 7 – 9 allmählich ändern. Es wird angenommen, dass die japanischen Versuchspersonen empfindlicher auf die fallenden Grundfrequenzkonturen als auf die steigenden reagieren, weil die fallende Grundfrequenzkontur im Japanischen bei der Wahrnehmung des Akzentes eine sehr wichtige Rolle spielt.
5. Wenn die gesamte Grundfrequenzkontur des Wortes "ata" nur steigend ist, halten mehrere japanische Versuchspersonen die zweite Silbe für akzentuiert. Der Grund dafür ist, dass diese Art von Akzent im Japanischen als 'unakzentuiert' oder als 'am Wortende akzentuiert' betrachtet wird.⁴ Im Fall eines unakzentuierten Wortes gibt es im Japanischen keine fallende Grundfrequenzkontur. Im Fall 'am Wortende akzentuiert' folgt am Ende eines Wortes ein funktionales Wort, das eine fallende Grundfrequenzkontur hat. In beiden Fällen hat die letzte Silbe eine hohe Grundfrequenz, und es gibt im ganzen Wort keine fallende Kontur.

4. Zusammenfassung und Schlussbemerkungen

Aufgrund der Ergebnisse dieser Arbeit wird angenommen, dass die japanischen Versuchspersonen bei der Wahrnehmung eines fremden Wortes entweder instinktiv oder bewußt auf ihre eigene Muttersprache Bezug nehmen. In dieser Arbeit wird nicht berücksichtigt, nach welchen Kriterien die Versuchspersonen die akzentuierten Silben markiert haben. Es kann auf die Ergebnisse Einfluss haben, ob sie die

Stimuli als Muttersprache oder als Fremdsprache gehört haben. In künftigen Arbeiten muss auch dies berücksichtigt werden.

Wenn es größere Datenmengen der europäischen Versuchspersonen gibt, wird der Vergleich der Ergebnisse zwischen mehreren Muttersprachlern sinnvoller.

*Frau Dawn M. Behne an der Technischen Universität Norwegen hat uns nützliche Anweisungen gegeben. Dafür möchten wir ihr danken. Ebenso danken wir den Studenten der Sophia-Universität für ihre Mitarbeit. Herrn Prof. Dr. Heinz Morioka möchten wir vor allem für seine geduldige Anweisungen bei der Beschreibung der phonetischen Arbeiten herzlich danken.

Anmerkungen

¹ Die Länge wird stets fixiert, um den Einfluss auf die Wahrnehmung des Akzentes zu vermeiden. Aus demselben Grund bleibt die Intensität des originalen Stimulus unberührt.

² Die Grundfrequenz in der vorderen Hälfte der ersten Silbe wird immer, wie bei der originalen Äußerung, um 130Hz fixiert.

³ Das Wort 'kategorial' wird hier nach dem englischen Terminus "categorical perception" benutzt.

⁴ Der Akzent des Tokyo-Japanischen basiert auf der Tonhöhe. Die Akzentmuster können in zwei Haupttypen eingeteilt werden: mit Akzentkern und ohne. Bei Wörtern mit Akzentkern sinkt die Tonhöhe unmittelbar nach dem Akzentkern plötzlich ab. Wenn es keinen Akzentkern gibt, bleibt die Tonhöhe gleich und es gibt keine plötzliche Senkung der Tonhöhe (flacher Akzenttyp). Die Wörter mit Akzentkern können nach der Position des Akzentkerns weiter in drei Akzenttypen klassifiziert werden: 1. Wenn die erste Mora akzentuiert wird, d.h. wenn eine fallende Grundfrequenzkontur auf die erste Mora folgt, wird das Wort 'Initial-hoch Akzenttyp (Atamadaka-Gata)' genannt. 2. Wenn der Akzentkern nicht auf der ersten und der letzten Mora eines Wortes, sondern auf einer der mittleren Moren liegt, wird das Wort 'Zentral-hoch Akzenttyp (Nakadaka-Gata)' genannt. 3. Wenn der Akzentkern auf der letzten Mora des Wortes liegt, und wenn die Grundfrequenz einer Partikel, die an das Wort anschließt, wie z.B. 'ga', plötzlich sinkt, wird das Wort 'Final-hoch Akzenttyp (Odaka-Gata)' genannt. (Saitou 1997)

Literatur

- Beckman, Mary E. 1986. *Stress and Non-Stress Accent*. Dordrecht: Foris Publications.
- Hayashi, R., Ayusawa T. und Nishinuma, Y. 1997. "Doitsugo Bogowasha-niyoru Tokyo-go Akusento-no Kikitori-Tesuto no Bunseki-Kekka", "21-Seiki-no-Nihongo-Onsei-Kyouiku-ni-mukete", Mombu-shou Kagaku-Kenkyuu-Hi "Kokusai-Shakai-ni-okeru-Nihongo-nitsuite-no Sougou-teki Kenkyuu." Kenkyuu-Daihyousha Mizutani, O., Kenkyuu-Han 3 "Onseigengo-no Inritsu-Tokuchou-ni-kansuru Jikken-teki Kenkyuu" *Houkokusho*, Kokuritsu-Kokugo-Kenkyuu-Jo: 31-36.
- Hirst, Daniel and Di Cristo, Albert. 1988. *Intonation Systems. A survey of Twenty languages*. Cambridge: Cambridge University Press.
- Masaki, Akiko, Takasawa, Miyuki, and Arai, Takayuki. 2000. "Go-akusento-no-Chikaku-niokeru Bogo-no-Eikyuu-ni-tsuite". *Proceedings of the 14th General Meeting of the Phonetic Society of Japan*: 165-170.
- Saitou, Yoshio. 1997. *Nihongo Onseigaku Nyuumon*. Tokyo: Sanseidou.
- Sugitou Miyoko. 1990. "Nihongo-to-Eigo-no Akusento-to-Intoneishon." *Kouza Nihongo-to-Nihongokyouiku*, ed. Miyoko Sugitou, 349-378. Tokyo: Meiji-shoin.